Abstract

Concept learning or learning from examples is one of the most extensively studied forms of learning in Artificial Intelligence(AI) and has applications in areas like Expert Systems, Pattern Recognition, etc The problem of concept learning involves learning to classify the objects of a domain using a set of pre-classified examples from the domain After seeing a set of such examples, the goal is to infer a concept description which can be used to classify new objects from the domain Because the problem involves inductive inference, the issue of validating the generalizations made during the learning process is important The Probably Approximately Correct(PAC) learning framework offers an elegant way to rigorously tackle this issue The PAC formulation is also useful to formulate the problem of learning in the presence of noisy examples Further, the framework serves as a unified formalism for analyzing and comparing complexity of learning algorithms

In this thesis, we address the problem of developing efficient incremental algorithms for concept learning from noisy examples under the PAC framework We employ logic expressions involving nominal and linear attributes to represent the concepts learnt In Chapter 2, we formulate an optimization problem that is relevant to concept learning and propose a general stochastic algorithm based on a novel team of Learning Automata to solve this problem Chapters 3 and 4 present applications of this optimization algorithm in learning logic expressions incrementally under noise The proposed algorithms are proved to converge asymptotically to the correct concept assuming that the examples given are corrupted by at most 50% of unbiased noise These algorithms can learn rich classes of logic expressions over both nominal and linear attributes Analyzing the time complexity of these learning algorithms is a related problem of interest This involves the problem of finite time analysis of Learning Automata algorithms, which is partially addressed in Chapter 5

Stochastic Optimization with a Hybrid Team of Learning Automata

Following Haussler's generalized PAC framework, concept learning involves an associated optimization problem of minimizing the empirical risk. In this thesis, we consider a stochastic iterative method to solve such optimization problems. We pose the optimization problem as a multi person game played by a team of learning automata. The game is novel in that each player of the game, represented by a Learning Automaton, can have a finite or an infinite action set to choose from. Defining solutions for the game called *optimal points*, we propose a decentralized learning algorithm. It is proved that the algorithm converges to one of the optimal points. We show that our result generalizes many of the currently available results in learning automata literature. An appealing aspect of our game formulation is that it can be employed as a general model for concept learning from noisy examples and our algorithm can be used for incrementally learning concepts as logic expressions. The convergence result proved can be used to show the correctness of the algorithm under PAC framework. We illustrate this application aspect of our game formulation next, by proposing algorithms for learning conjunctive and disjunctive logic expressions.

Learning Conjunctive Concepts

We first consider the problem of learning conjunctive concepts. In the literature, algorithms for efficiently learning conjunctive concepts over nominal and linear attributes and possessing PAC learnability properties have been developed under no noise conditions. However, extending these algorithms to handle noisy examples is difficult. Nonincremental algorithms that tackle either nominal attributes only or linear attributes only, have been proposed to learn from noisy examples. Though a blend of these algorithms could be used to learn concepts with nominal and linear attributes, the resulting algorithms would still be nonincremental. Incremental learning algorithms are more efficient, especially in problems where large number of training examples are needed. In this thesis, we consider the problem of incremental learning of a class of conjunctive concepts involving nominal and linear attributes, called *simple conjunctive concepts*, from noisy examples. We propose an algorithm using the hybrid automata team model considered above and prove that the algorithm correctly learns the class of simple conjunctive concepts under up to 50% noise. Through simulations on some synthetic and real-world problems, we show that the algorithm is computationally efficient also.

Learning Disjunctive Concepts with Marker Attributes

We next consider the problem of learning disjunctive concepts In view of the rich representational capabilities of disjunctive concepts, the problem has been well-investigated in the literature However, the algorithms available for provably learning under PAC framework are very few, especially under noise Even under no noise, the available algorithms for learning in the presence of nominal and linear attributes turn out to be nonincremental In this thesis, we observe that learning of disjunctive concepts is inherently difficult because of a credit assignment problem We provide an approach to resolve the credit assignment and hence efficiently learn disjunctive concepts, using special nominal attributes of the domain, called *marker attributes* Using the idea of marker attributes, we define concept classes called *k term disjunctive concepts with marker attributes* We show that these classes are efficiently learnable under no noise, with polynomial sample and time complexity In the presence of noise, we propose incremental learning algorithms based on hybrid automata team models We prove that all these algorithms correctly learn under the PAC framework It is also proved that the algorithms learn under up to 50% noise in the examples We also exhibit the computational efficiency of the algorithms through simulation studies on few synthetic and real world problems When the marker attributes are known *a priori*, all our algorithms have more efficient versions both under noise-free and noisy conditions

We observe that concepts with marker attributes form nontrivial subclasses of disjunctive concepts It is shown that these classes represent a reasonably large class of disjunctive concepts by varying the number of marker attributes We argue that the restriction of the class of all disjunctive concepts to concepts with marker attributes offers an interesting trade-off between expressive power and complexity of learning For example, by considering a sub class of k-term boolean expressions(over N variables/attributes) expressible with at most, say $(k/2)$, marker attributes, we show that this subclass is learnable by an algorithm that is more efficient by $O(N^{\frac{k}{2}-1})$ factor compared to the algorithm of Pitt & Valiant

Finite time Behaviour of Learning Automata

After having seen that algorithms based on teams of learning automata are effective for learning concepts under the PAC framework, we illustrate how PAC framework offers possibilities for analyzing automata algorithms Such an analysis would be useful in deriving theoretical estimates on the learning complexity of algorithms based on learning automata models, presented above We propose a general framework for analyzing the finite time behaviour of the automaton learning algorithms motivated by the PAC formulation Using

this framework, the finite time analysis of the Pursuit Algorithm is presented We consider both continuous and discretized forms of the pursuit algorithm Based on the results of the analysis, we compare the rates of convergence of these two versions of the pursuit algorithm

To conclude, we propose in this thesis, a novel class algorithms based on Learning Automata models for incremental concept learning from noisy examples under the PAC framework Defining a new class of disjunctive logic expressions involving nominal and linear attributes we show that the algorithms can efficiently learn these rich subclasses of disjunctive concepts The algorithms are provably noise-tolerant and can be run in an on-line fashion They are parallel algorithms and can be easily implemented on SIMD machines with processors doing only local computation With such implementation, we can expect close to linear speedup of our algorithms